

Sensory Architectures for Biologically Inspired Autonomous Robotics

CHARLES M. HIGGINS

Neuromorphic Vision and Robotic Systems Laboratory, Department of Electrical and Computer Engineering, The University of Arizona, 1230 E. Speedway Blvd., Tucson, Arizona 85721

Abstract. Engineers have a lot to gain from studying biology. The study of biological neural systems alone provides numerous examples of computational systems that are far more complex than any man-made system and perform real-time sensory and motor tasks in a manner that humbles the most advanced artificial systems. Despite the evolutionary genesis of these systems and the vast apparent differences between species, there are common design strategies employed by biological systems that span taxa, and engineers would do well to emulate these strategies. However, biologically-inspired computational architectures, which are continuous-time and parallel in nature, do not map well onto conventional processors, which are discrete-time and serial in operation. Rather, an implementation technology that is capable of directly realizing the layered parallel structure and nonlinear elements employed by neurobiology is required for power- and space-efficient implementation. Custom neuromorphic hardware meets these criteria and yields low-power dedicated sensory systems that are small, light, and ideal for autonomous robot applications. As examples of how this technology is applied, this article describes both a low-level neuromorphic hardware emulation of an elementary visual motion detector, and a large-scale, system-level spatial motion integration system.

Engineers Have a Lot to Gain From Studying Biology

The architecture of biological sensory systems has something to teach conventional engineering about system-level design. Neurobiological representations of visual modalities including depth, motion, color, and form are quite unlike those employed by conventional computer vision systems. Neurobiological computational architectures are not modular, linear, or feedforward. And yet biological organisms routinely accomplish complex visual tasks such as object recognition, obstacle avoidance, and target tracking, which continue to challenge artificial systems.

Dealing with the complexity posed by the availability of a multitude of parallel sensors and a large number of redundant actuators is a major unsolved problem in modern computational systems. As we become more ambitious in our inclusion of computational systems into every possible device, more sensory inputs are available; more actuators are controlled by the computing system; more possible conditions must be recognized and appropriately dealt with. Engineers in fields from space-bound robotics and guided missile design to automotive navigation and prosthetics are beginning to realize the value of paying attention to biological solutions. The biological literature is replete with detailed analyses of systems that deal with complex sensory and motor control tasks and still manage to outperform, by any metric, anything created by mankind. In the papers in this collection (CASSLS, 2001), for example, Robert Barlow and colleagues describe a finely elaborated model of horseshoe crab visual processing, revealing how the crab manages to navigate and find mates in a highly variable and noisy underwater environment (Barlow *et al.*, 2001); Thomas Cronin and Justin Marshall give us insight into the marvelously complex mantis shrimp visual system, a

E-mail: higgins@ece.arizona.edu

This paper was originally presented at a workshop titled *Invertebrate Sensory Information Processing: Implications for Biologically Inspired Autonomous Systems*. The workshop, which was held at the J. Erik Jonsson Center for the National Academy of Sciences, Woods Hole, Massachusetts, from 15–17 April 2000, was sponsored by the Center for Advanced Studies in the Space Life Sciences at the Marine Biological Laboratory, and funded by the National Aeronautics and Space Administration under Cooperative Agreement NCC 2-896.

unique and highly specialized tool for predation (Cronin and Marshall, 2001); Mandyam Srinivasan and co-workers introduce us to the secrets of insect vision (Srinivasan *et al.*, 2001).

The daunting task for engineers is to produce—by abstraction from the specific organism—a set of general design principles to replace the top-down, linear, functional-block oriented strategies that characterize conventional design. These principles may include layered massively parallel processing, population coding, probabilistic information transmission, a mix of analog and digital coding strategies, connection-oriented architectures, removal of redundant information at every processing stage, positive as well as negative feedback, and the use of nonlinear computational elements (such as voltage-dependent conductances and neurons with adaptive properties), even though these elements make the system more difficult to describe analytically in closed form. Although the mathematical principles underlying conventional design are well known, it is not clear that we have the mathematical tools to analyze and design systems such as those suggested by neurobiology. However, these tools are being developed in the theory of nonlinear dynamical systems, as well as in the study of neurobiological systems themselves.

By building artificial sensory systems true to the representations and computational architectures used by neurobiology, we will be able to produce novel, highly capable autonomous robots, while at the same time suggesting testable hypotheses as to how biological systems accomplish sensory and motor tasks.

Efficient Implementation of Biological Computational Architectures

Because the primary computational strategy used by neurobiology is layered, massively parallel processing, an efficient parallel implementation is essential to make effective use of such computational architectures. Biological computational architectures, which are continuous-time parallel algorithms, map very badly onto conventional discrete-time serial processors. The conventional real-time implementation of a biological vision computation involves an imager providing discrete-time frames to a high-speed signal processor that *serially* performs a large number of operations that could be performed in *parallel*. In contrast, true parallel architectures may be straightforwardly implemented in neuromorphic VLSI (very large scale integration) hardware (Mead, 1989). This type of special-purpose hardware implements biological algorithms directly and efficiently in patterns of transistors, resistors, and capacitors. By making use of the physical primitives provided by electronic de-

vices, highly efficient computations can be performed on signals represented by currents and voltages. Neuromorphic technology focuses on analog computation for efficiency, but in general mixes both analog and digital transistor circuitry. The implementation can be at a level of abstraction from the biophysical (Hahnloser *et al.*, 2000) to the neural (Boahen, 1999) to the mathematical (Higgins and Korrapati, 2000). The individual physical components in such a design are on the scale of micrometers and are fabricated together on a single piece of silicon measuring millimeters on a side, resulting in a physically robust compact package.

While general-purpose digital processors grow ever more powerful, special-purpose hardware will always be more efficient in power consumption and size for specific problems. This is because the former, by its nature, is not optimized for any specific computation, whereas the latter includes only the circuitry that is necessary for the problem at hand. The efficiency of a task-specific analog implementation comes at the price of precision and flexibility. The analog implementation of a computation is subject to inevitable noise that limits the possible precision, whereas a digital implementation can (potentially but usually not practically) be performed with virtually unlimited precision. The software implementation of a computation clearly has flexibility to change at the programmer's whim, whereas a dedicated hardware implementation is fixed at fabrication time. With awareness of these trade-offs, a designer will choose a special-purpose hardware solution where the power and size constraints are most important. This is often an appropriate choice in real-time vision applications on autonomous robots.

An Example of Neuromorphic Implementation: Low-Level Motion Detection

The neurobiological representation of visual motion in organisms from insects to primates is in terms of oriented spatiotemporal-frequency-tuned cells. The conventional *optical flow* representation is mathematically a vector field, allowing only one velocity of motion in every local area of the visual scene. The biological representation of visual motion is more powerful in that it allows multiple simultaneous directions of motion as long as they are distinguished in spatial or temporal frequency, making possible our perception of transparent motion and occlusion boundaries. However, the biological representation of motion requires a range of spatial- and temporal-frequency tunings to cover all stimuli of interest. Primates have a tremendous number of neurons (on the order of 10^{11}) and thus can afford to populate the whole spatiotemporal frequency space with a

large bank of tuned neurons for each spatial location. Insects, with a much smaller set of neurons (on the order of 10^5), are forced to use a much smaller set of tunings, yet they operate very effectively in the same environment. Thus insects may provide engineers with a more neuron-efficient example of a visual motion system.

In a seminal 1985 paper, Adelson and Bergen proposed a model of the motion response of primate complex cells and showed it to be equivalent to the Reichardt (1961) model of elementary motion detection in the fly. We have implemented this *motion energy* model in neuromorphic VLSI hardware (Higgins and Korrapati, 2000), both to allow system-level hardware modeling of biological motion systems, addressed in the next section, and to provide a sensitive, real-time visual motion detector for robotic applications. Because of the lack of any digital thresholding step in the algorithm, motion energy sensors (like the motion-sensitive cells they model) are limited in contrast response only by signal-to-noise ratio, allowing them to perceive the direction of motion even for very low contrasts by integrating over time. The original algorithm, shown in Figure 1A, combines the response of four simple cells with a rectifying nonlinearity (a square) to create a spatiotemporal-frequency-tuned estimate of motion direction. Figure 1B shows how this model was implemented in hardware: several simplifications were made in the pursuit of a compact implementation, but the algorithm still performs a highly sensitive directional motion computation, as shown in Figure 2. This sensor was implemented with only 41 transistors, allowing hundreds or thousands of parallel sensors on a reasonably sized silicon substrate. Because of the analog implementation, each sensor consumes less than 40 microwatts of power, allowing large-scale parallel implementation without problematic power consumption or special cooling needs.

System-Level Neuromorphic Design: Spatial Motion Integration

As mentioned earlier, the spatiotemporal filter representation of motion common to a wide variety of organisms requires a range of spatial- and temporal-frequency tunings to cover all stimuli of interest and thus, unlike the conventional representation, must be implemented with a bank of filters for each spatial location. Once this bank of filters is computed, the information must be spatially integrated to be of use to the organism; in other words, information about spatial *patterns* of visual motion is more useful to the animal than the low-level motion information itself. This integration occurs in insect tangential neurons (Krapp and Hengstenberg, 1996) and in the medial superior temporal area (MST) of primate visual cortex (Perrone and Stone,

1998). This strategy suggests a computational architecture like that shown in Figure 3A, which includes a first-stage photosensitive array, a second stage of multiple parallel motion-processor arrays computing different spatial frequencies, orientations, and temporal frequencies, and a third integration stage that can synthesize units sensitive to wide-field spatial patterns of motion by combining spatial regions of low-level motion detectors.

We have implemented a preliminary version of such an architecture in neuromorphic VLSI hardware (Higgins and Shams, unpubl. data), as shown in Figure 3B. Each two-dimensional processing unit is implemented as a separate VLSI chip; chips communicate using trains of spikes modeling action potentials. This multi-chip neuromorphic VLSI system is capable of synthesizing units sensitive to complex arbitrary patterns of visual motion, including expansion, contraction, rotation, and translation, as shown in Figure 4. These patterns, along with more elaborate ones, allow determination of imager self-motion through the world for postural stability and navigation (Zemel and Sejnowski, 1998), tracking of objects (Gronenberg and Strausfeld, 1991), and obstacle avoidance (Gabbiani *et al.*, 1999). Despite the multi-chip nature of the system, it is still a compact, robust, low-power special-purpose computing system ideal for robotic applications.

Summary

This article contends that biological representations and computational architectures are useful in a wide range of engineering problems. However, in practice, it is difficult to separate biological algorithms from their implementation, and so we also present an implementation technology, neuromorphic VLSI, which is well suited for physical realization of these algorithms. A low-level motion processor based on the Adelson-Bergen model was described to illustrate how a biological model may be directly mapped onto an analog VLSI transistor circuit for highly efficient implementation. From an engineer's perspective, this sensor benefits not only from the biological representation, but also from the analog nature of the biological model computation. A multi-stage motion-processing system was also described, exemplifying a biologically inspired computing architecture that also uses representations inspired by biology. This architecture is very well matched to the vision problem and allows a specific useful computation to be done with very low requirements for power consumption and size. It is the author's contention that systems of this type are the best hope for future highly capable small autonomous robotic systems.

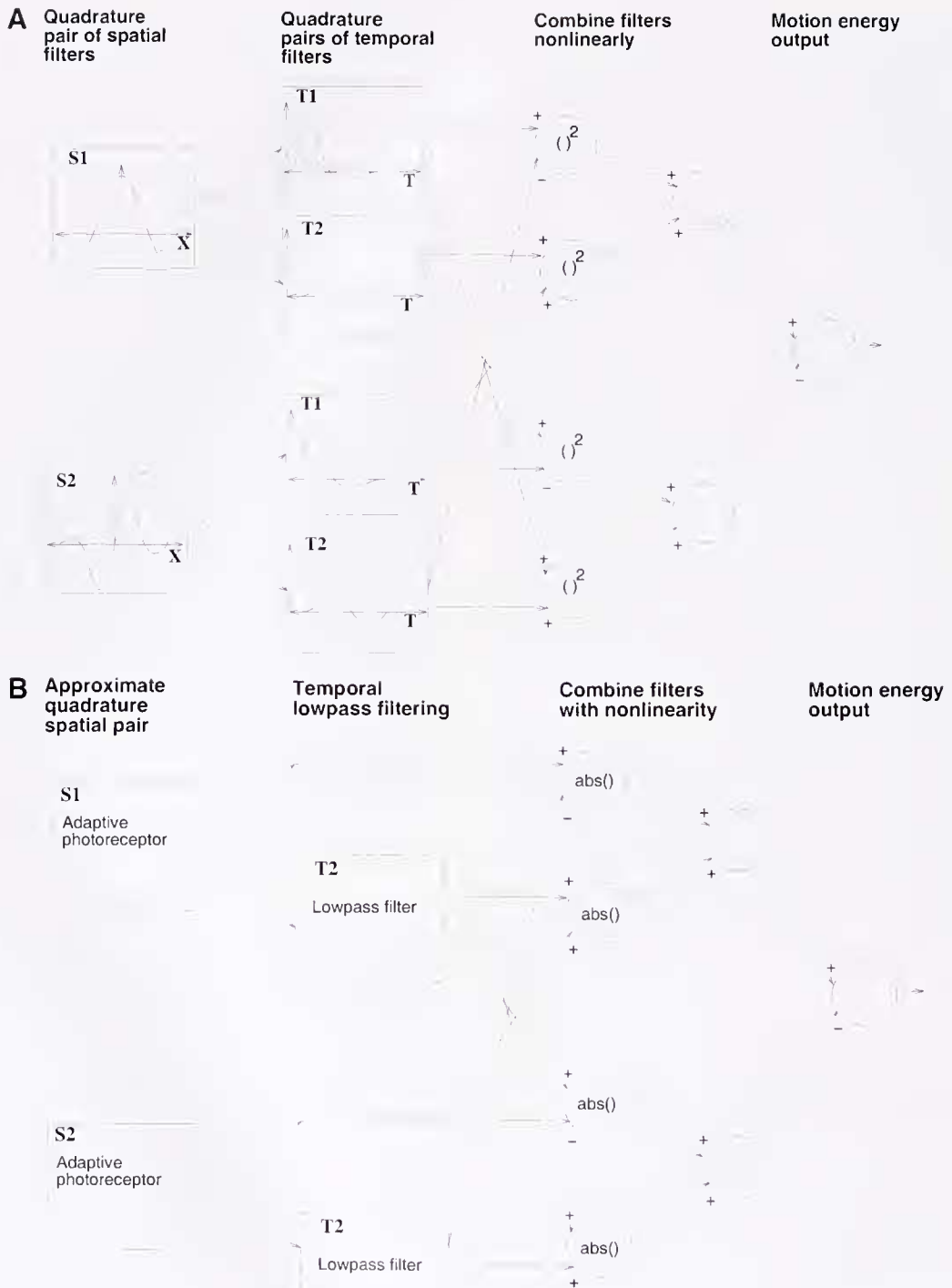


Figure 1. Low-level motion detection. (A) The canonical Adelson-Bergen motion energy model combines each of an even and odd spatial Gabor filter with two temporal handpass filters to create the output of four linear simple cells. The output of these simple cells is combined nonlinearly in quadrature pairs to create a phase-independent motion direction estimate, which models a primate complex cell. (B) For hardware implementation, the spatial filters were implemented with a center-surround diffusive network, and the temporal filters by using a single lowpass filter. Rather than using a squaring operation, a simpler rectifying nonlinearity (the absolute value) was employed. The output of the implementation is quite comparable to the original model.

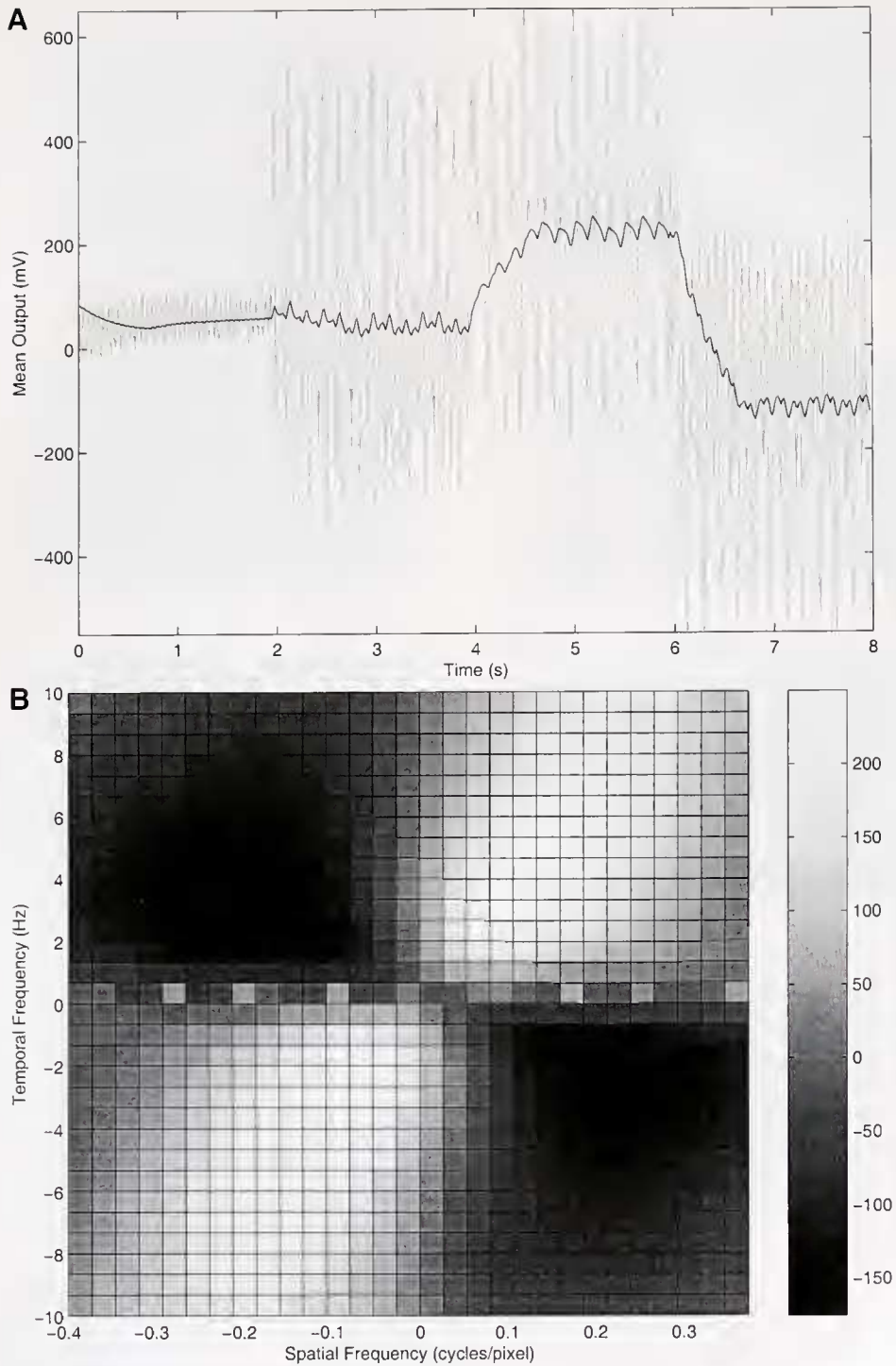


Figure 2. Output of the low-level motion detector. (A) The output of the hardware motion sensor is shown in response to no stimulus (under fluorescent lighting), an orthogonal, preferred and null direction sinusoid grating. The lighter traces represent raw output; darker lines are averaged output. (B) As the spatial and temporal frequency of a moving sinusoidal stimulus is varied, the sensor shows a strong preference in a certain band of frequencies.

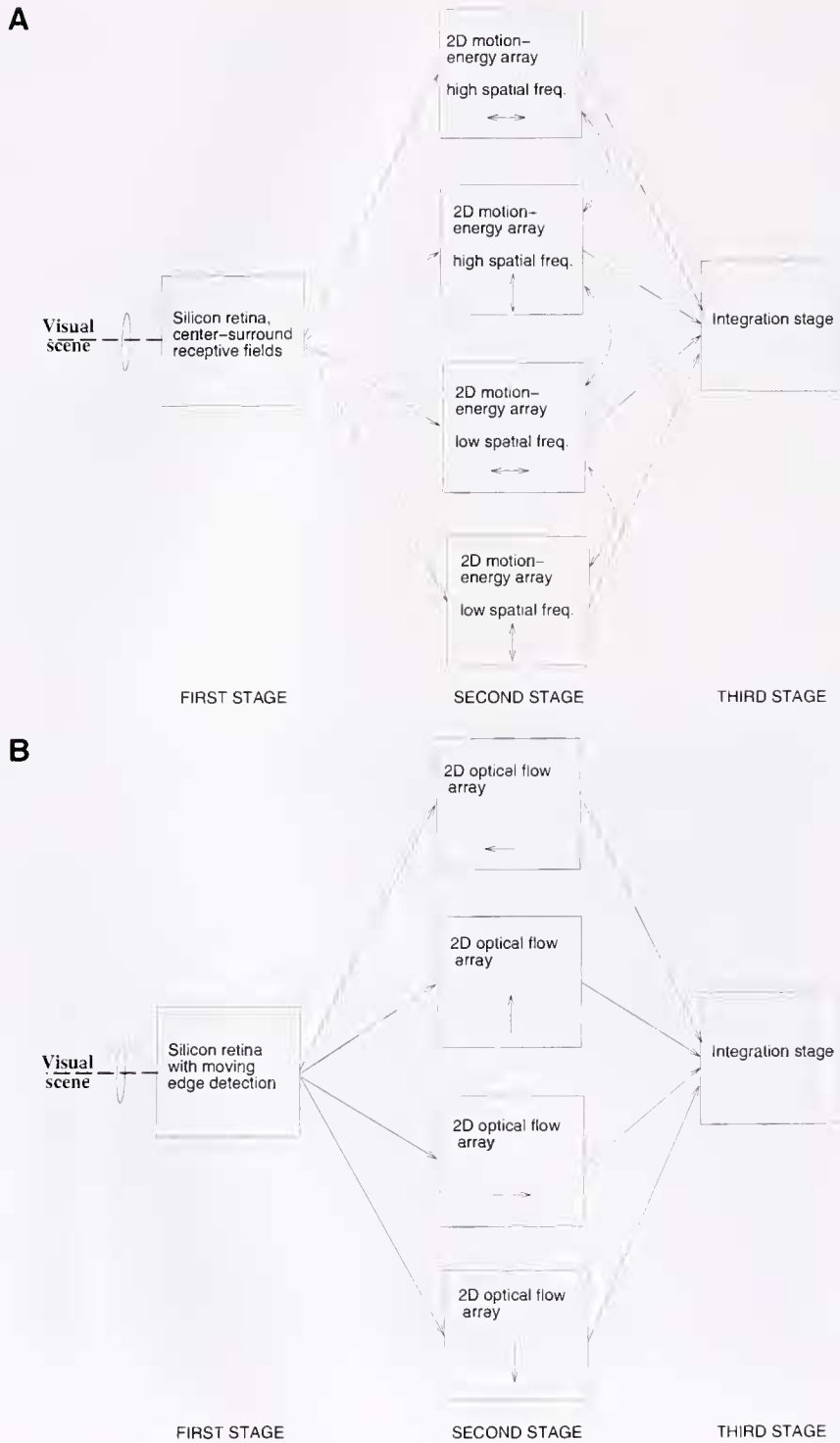


Figure 3. Architecture for spatial motion integration. (A) In the idealized system, a first-stage silicon retina is followed by multiple second-stage motion processors tuned for different orientations and spatial and temporal frequencies. Interaction is allowed between second-stage motion processors. The final stage integrates over space to synthesize sensitivity to patterns of visual motion. (B) In the current system, the silicon retina detects moving edges and transmits this information to the next stage. The second-stage motion processors use the moving edge information to compute the direction of edge motion, with each of four processors tuned for a different direction of motion. The final stage synthesizes spatial motion sensitivity.

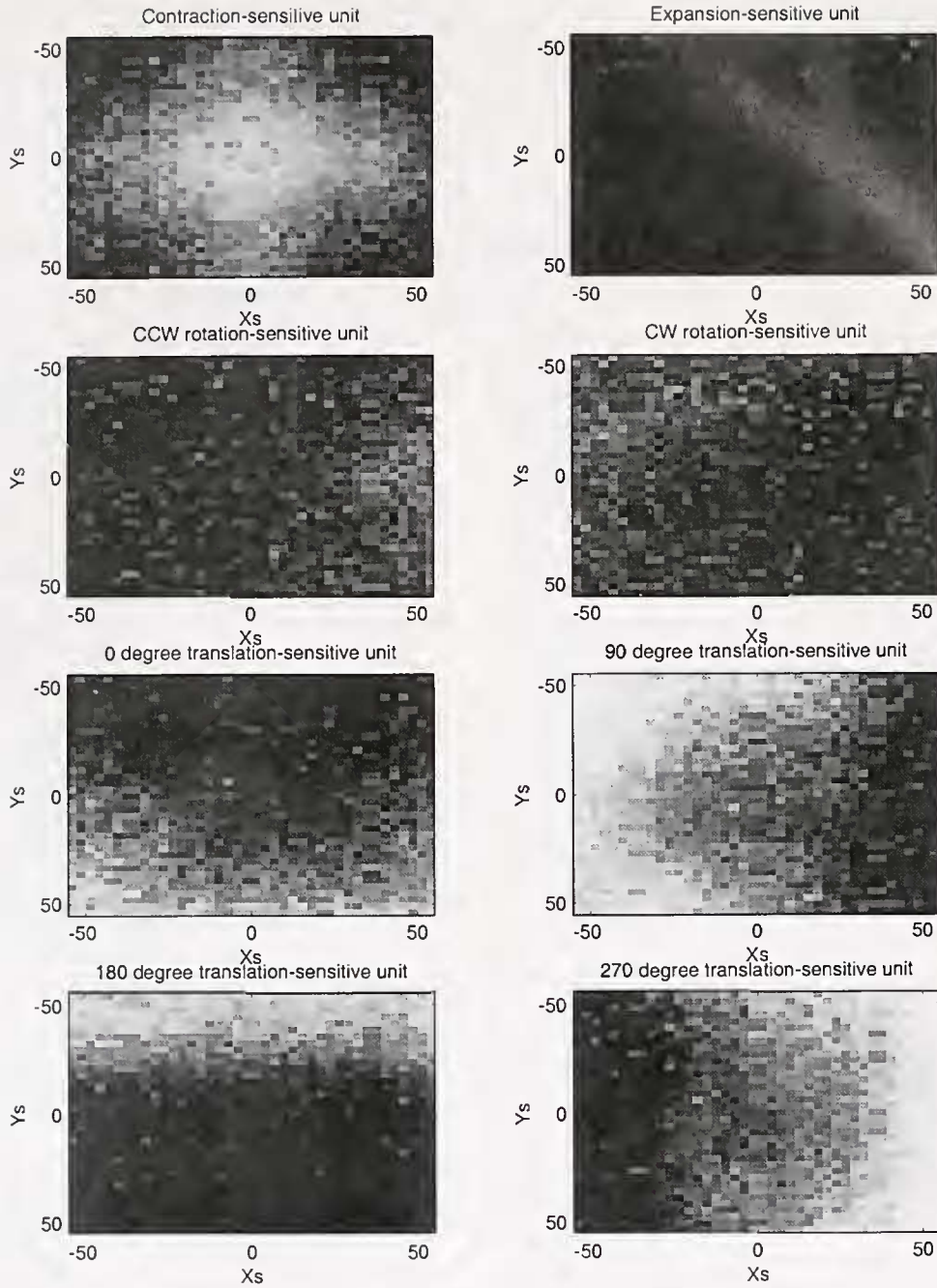


Figure 4. Performance of the spatial motion integration system. This figure shows the response of eight simultaneously synthesized units to a wide-field stimulus—a periodically contracting circle—as the focus of contraction is moved around the visual field. Lighter shading indicate stronger responses. The upper left unit is tuned for contraction at the center of the visual field and thus shows a strong response there. The upper right unit is tuned for expansion and thus shows no response. In the second row from the top, units tuned for clockwise and counterclockwise rotation likewise show little or no response. However, units shown at the bottom tuned for translation in four different directions show a strong response when the focus of contraction nears the appropriate edge of the visual field, because the farther the contracting pattern gets from the center, the more similar it becomes to a simple translating pattern.

Acknowledgments

The author gratefully acknowledges support for this research by the National Science Foundation through the California Institute of Technology Center for Neuromorphic Systems Engineering as part of the Engineering Research Centers program, by the Office of Naval Research under grant number N00014-97-1-0970, and by the National Science Foundation's State/Industry/University Cooperative Research Center for Low Power Electronics (CLPE). CLPE is supported by NSF (Grant #EEC-9523338), the State of Arizona, and the following companies and foundations: Burr-Brown, Inc., Conexant, Gain Technology, Intel Corporation, Medtronic Microelectronics Center, Microchip Technology, Motorola, Inc., The Motorola Foundation, Raytheon, Texas Instruments, and Western Design Center. The author also thanks the anonymous reviewers for their assistance in clarifying this manuscript.

Literature Cited

- Adelson, E. H., and J. R. Bergen. 1985. Spatiotemporal energy models for the perception of motion. *J. Opt. Soc. Am. A* 2(2): 284-299.
- Barlow, R. B., J. M. Hitt, and F. Dodge. 2001. *Limulus* vision in the marine environment. *Biol. Bull.* 200: 169-176.
- Boahen, K. 1999. Retinomorphing chips that see quadruple images. In *Proceedings of the 7th International Conference on Microelectronics for Neural, Fuzzy and Bio-inspired Systems*, University of Granada, Spain, 7-9 April 1999.
- CASSLS (The Center for Advanced Studies in the Space Life Sciences). 2001. Invertebrate sensory information processing: implications for biologically inspired autonomous systems. *Biol. Bull.* 200: 145-245.
- Cronin, T., and J. Marshall. 2001. Parallel processing and image analysis in the eyes of mantis shrimps. *Biol. Bull.* 200: 177-183.
- Gabbiani, F., H. Krapp, and G. Laurent. 1999. Computation of object approach by a wide-field, visual neuron. *J. Neurosci.* 19: 1122-1141.
- Gronenberg, W., and N. J. Strausfeld. 1991. Descending pathways connecting the male-specific visual system of flies to the neck and flight motor. *J. Comp. Physiol. A* 169: 413-426.
- Hahnloser, R. H. R., R. Sarpeshkar, M. A. Mahowald, R. J. Douglas, and H. S. Seung. 2000. Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit. *Nature* 405: 947-951.
- Higgins, C. M., and S. K. Korrapati. 2000. An analog VLSI motion energy sensor based on the Adelson-Bergen algorithm. In *Proceedings of the International ICSC Symposium on Biologically-Inspired Systems*, Wollongong, Australia, 12-15 December 2000.
- Krapp, H. G., and R. Hengstenberg. 1996. Estimation of self-motion by optic flow processing in single visual interneurons. *Nature* 384: 463-466.
- Mead, C. A. 1989. *Analog VLSI and Neural Systems*. Addison-Wesley, Reading, MA.
- Perrone, J. A., and L. S. Stone. 1998. Emulating the visual receptive-field properties of MST neurons with a template model of heading estimation. *J. Neurosci.* 18(15): 5958-5975.
- Reichardt, W. 1961. Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. Pp. 303-317 in *Sensory Communication*, W. A. Rosenblith, ed. MIT Press, Cambridge, MA.
- Srinivasan, M. S. Zhang, and J. S. Chahl. 2001. Landing strategies in honeybees, and possible applications to autonomous airborne vehicles. *Biol. Bull.* 200: 216-221.
- Zemel, R. S., and T. J. Sejnowski. 1998. A model for encoding multiple object motions and self-motion in area MST of primate visual cortex. *J. Neurosci.* 18(1): 531-547.